



HAL
open science

A Study of Element Ordering in English Coordinate Lexical Items

Vincent Renner

► **To cite this version:**

Vincent Renner. A Study of Element Ordering in English Coordinate Lexical Items. *English Studies*, 2014, 95 (4), pp.441 - 458. hal-01097876

HAL Id: hal-01097876

<https://hal.science/hal-01097876v1>

Submitted on 22 Dec 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Study of Element Ordering in English Coordinate Lexical Items

Vincent Renner

This article aims at determining the statistical validity of various constraints which are held to influence the element ordering of coordinate lexical items in English. Twelve constraints mentioned in the literature are tested on 562 binomials, compounds and blends. Nine of them are found to be statistically significant: experiential closeness, temporal iconicity, syllable number, vowel length, initial consonant obstruency, final consonant obstruency, stress alternation and lexical frequency for binomials; temporal iconicity and syllable number for compounds; syllable number and initial consonant complexity for blends.

1. Introduction

Lexical items composed of semantically coordinate elements are original units of the lexicon in so far as their internal ordering is not as strictly constrained as in the case of subordinate elements. The ordering is generally somewhat arbitrary until the structure is institutionalized, and even then, it may happen that the two orderings are listed as synonyms (e.g. *doom and gloom / gloom and doom, lend-lease / lease-lend, ruddevator / elerudder*). Arbitrariness is not absolute as element ordering is held to be influenced by a variety of constraints, mainly of semantic and phonological nature. The constraints mentioned in the literature are, however, usually tested only on a relatively small number of items, and they are not supported by statistical evidence, which raises doubts on their actual weight. The aim of the present study is to provide substantiated insight into these lexical structures by measuring the statistical validity of twelve constraints listed in the literature. The constraints will be tested on the whole class of English coordinate lexical items, which comprises binomials (e.g. *bed and breakfast, cut and paste, nickel-and-dime, trick-or-treat*), compounds (e.g. *deaf-mute, gum resin, hunter-gatherer, tractor-trailer*) and blends (e.g. *cafetorium, fantabulous, modem, tangelo*).

2. Literature Review

2.1 Semantic Constraints

A detailed study of the semantic constraints governing element ordering in English binomials is provided by William Cooper and John Ross.¹ A list of twenty-two dominant notions is given:

- (1) Here (*here and there, in and out*)
- (2) Now (*now and then, sooner or later*)
- (3) Present Generation (*father and grandfather, son and grandson*)
- (4) Adult (*father and son, mare and foal*)
- (5) Male (*husband and wife, brother and sister*)
- (6) Positive (*more or less, win or lose*)
- (7) Singular (*singular and plural, once or twice*)
- (8) Patriotic (*cowboys and Indians, United States and Canada*)
- (9) Animate (*people and things, men and machines*)
- (10) Friendly (*friend or foe, pro and con*)
- (11) Solid (*land and sea, Army and Navy*)
- (12) Front (*front and back, fore and aft*)
- (13) Agentive (*speaker and hearer, cat and mouse*)
- (14) Power Source (*bow and arrow, horse and carriage, gin and tonic*)
- (15) Living (*life and death, live or die*)
- (16) At Home (*at home and abroad, home and away*)
- (17) General (*general and particular, abstract and concrete*)
- (18) Nominal (*nouns and verbs*)
- (19) Count (*count and mass nouns*)
- (20) Food and Drink Hierarchy: Fish > Meat > Drink > Fruit > Vegetables > Baked Goods > Dairy Products > Spices (*fish and game, meat and drink, surf and turf, fish and chips, peaches and cream*)
- (21) Divine (*God and man, church and state*)
- (22) Plant (*plant and animal, flora and fauna*)

¹ Cooper and Ross, 64–7.

Cooper and Ross admit that exceptions exist (e.g. *past and present*, *ladies and gentlemen*, *dead or alive*), but they stress that binomials overwhelmingly respect the constraints. Their analysis is on the whole insightful, but if the notions and illustrations given are considered in detail, the classification does not appear to be completely satisfactory. Some notions are not clearly delineated (e.g. Present Generation) while others seem to have been created unnaturally to group heterogeneous collections of binomials (Singular is used to unite *Mick Jagger and the Rolling Stones* and *monotheism and polytheism*, Power Source for *bow and arrow* and *gin and tonic*, General for *form and substance* and *word and deed*). Several notions also have a very limited scope – (3), (18), (19) and (22) are illustrated by only one or two binomials – and others could easily be merged (Friendly could be included in Positive; Animate and Living are quasi-synonyms; Here, Patriotic and At Home could be merged into Spatial Closeness, Now and Present Generation into Temporal Closeness).

Cooper and Ross consider that a general semantic principle can be abstracted and they group most of the constraints under an umbrella concept that they term “*Me First principle*”. The left position in a binomial is generally to be reserved for the element in empathy with a virtual prototypical speaker’s vision of the world. This principle is known under a variety of labels in the literature: John Ross facetiously uses the term *myopia* (“*The ego is at the center of the speaker’s conceptual space. It is focused inward, it is shortsighted [...].*”), Suitbert Ertel speaks of “*closeness to the speaker’s cognitive field*”, Susumu Kuno of “*speaker’s empathy*”, Keith Allan of “*familiarity hierarchy*”, Willy van Langendonck of “*closeness hierarchy*”.²

Two other constraints, which do not receive adequate treatment in Cooper and Ross’s study,³ can be traced back to Richard Abraham and Yakov Malkiel,⁴ and have reappeared in more recent literature. Jolanta Szpyra and Marge Landsberg both stress the prominent role played by Temporal Iconicity, a constraint requiring that the element which is first chronologically be first in the coordinate structure (the principle is typically illustrated by the famous quote from Julius Caesar, “*Veni, vidi, vici.*”, and explains the ordering in items such as *catch and release* and *cook-chill*).⁵ As for the second constraint, Szpyra speaks of the notion Principal, and gives the binomials *bed and board* and *fish and chips* as illustrations; Landsberg uses the phrase “*more important precedes less important*” to account for the internal ordering in the syntagma *the President and the Secretary of State*. Two linguists commenting on the internal ordering of French coordinate compounds have also remarked on the asymmetrical relation between the left element, which is more salient, and the

² Ross, “Myopia” (quoted in Birdsong, “Iconicity”, 40–1); Ertel (quoted in Fenk-Oczlon, 521); Kuno (quoted in Fenk-Oczlon, 521); Allan, 52–4; Van Langendonck, 83.

³ At the very end of their article, Cooper and Ross briefly acknowledge the existence of the constraint Chronology (e.g. *eat and run*, *stop and shop*).

⁴ Abraham, 284–5; Malkiel, 145–8.

⁵ Szpyra, 36, 39; Landsberg, 72.

right element, which is comparatively less salient: Michael Picone speaks of the “*secondariness*” of the right element in a compound such as *canapé-lit* (*sofa bed*), Pierre Arnaud of a “*salience hierarchy*” (Fr. *hiérarchie de prégnance*) between the denotata of the compounding elements in a compound such as *député-maire* (*MP-cum-mayor*).⁶

In the present study, only three broad notions – those which are clearly established – have been retained in the set of semantic constraints to be tested. First is Experiential Closeness, which corresponds to the unmarkedness of the most salient conceptual features – [POSITIVE], [HUMAN], [ANIMATE], [MALE] and [ADULT] – that Cooper and Ross associate with the virtual prototypical speaker in English; second is Superiority (both in spatial and hierarchical terms); third is Temporal Iconicity. For those notions which are more loosely associated with the Me First principle, an explanation making use of the notion of salience differential seems more appropriate than Cooper and Ross’s approach as it can be adapted to individual lexical items – *bow and arrow* and *gin and tonic* for instance do not need to be awkwardly grouped under Power Source anymore.

2.2 Phonological Constraints

The most detailed study of the phonological factors conditioning the internal ordering of binomials is again the work of Cooper and Ross.⁷ They list seven features which differentiate the left element from the right one. Other factors being equal, the right element contains:⁸

- (1) more syllables (*vim and vigor, wild and woolly*)
- (2) a longer vowel (*stress and strain, trick or treat*)
- (3) more initial consonants (*sink or swim, make or break*)
- (4) a more obstruent segment if the two elements have only one initial consonant⁹ (*wear and tear, surf and turf*)
- (5) a vowel with a lower second-formant frequency (*cats and dogs, this and that*)
- (6) fewer final consonants (*betwixt and between, wax and wane*)
- (7) a less obstruent segment if the two elements have only one final consonant¹⁰ (*push and pull, rock and roll*)

⁶ Picone, 125–7; Arnaud, 330.

⁷ Cooper and Ross, 71–82.

⁸ The constraints are presented by Cooper and Ross in a hierarchical order, from the most to the least potent.

⁹ Cooper and Ross use the following hierarchy of obstruency: stops > fricatives > nasals > liquids > glides.

¹⁰ Cooper and Ross propound the following hierarchy of vowels: [i] > [ɪ] > [e] > [æ] > [ɑ] > [ɔ] > [ou] > [u]. John Oakeshott-Taylor, who experimentally tested ordering preferences by using nonsensical binomials of

The original analysis was slightly modified by Steven Pinker and David Birdsong, who experimentally tested five of the seven constraints – (1), (2), (4), (5) and (6) – on the basis of the preferred ordering of nonsensical elements. They validated four out of the five constraints:

- (1) Syllable Number (*dack and badack*)
- (2) Vowel Length (*brets or braits*)
- (4) Initial Consonant Obstruency (*haipo and daipo*)
- (5) Vowel Quality¹¹ (*gligy and glagy*)

The last constraint – (6) Final Consonant Complexity – was invalidated and reversed: the element containing more final consonants will preferably appear in second position (e.g. *flar and flard*).¹² The volte-face is not problematic: the element ordering of the two illustrative binomials given by Cooper and Ross can still be accounted for, as *betwixt and between* and *wax and wane* both respect Vowel Length. This reversal is even remarkable in so far as (2), (3) and (6) obey a common principle: the three constraints favour the shorter element in initial position. Furthermore, if one considers that a more open and/or more back vowel is, acoustically speaking, longer, (2) and (5) can be merged.¹³ If, in addition, (1) is incorporated, the result is a general phonological principle – called “*short/long contrast*” by John Ross and “*final-heavy principle*” (Fr. *loi du second lourd*) by Claude Hagège¹⁴ – which demands that the structurally simpler element precede the more complex one.

One of the above constraints, Syllable Number, which is also known as Panini’s Law in the literature, has also been tested on blends. Michael Kelly has calculated that the source lexemes corresponding to the splinters occupying the first position in a blend average 2.2 syllables, and those corresponding to the splinters in second position 2.7 syllables.¹⁵ The discrepancy is confirmed by Stefan Gries’s statistics: out of his list of 1028 blends, 24.3% have equisyllabic source lexemes,

the pV_1t and pV_2t type, obtains a fairly similar hierarchy: [ɪ] > [i:] > [e] > [æ] > [ɒ] > [a:] > [ʊ] > [ɔ:] > [u:] > [ɜ:].

¹¹ Pinker and Birdsong (quoted in Birdsong, “Iconicity”, 33) substitute the concept of vowel quality to that of second-formant frequency and reformulate the constraint as follows: “*First-ordered constituents will have more closed or more front vowels, while second constituents will have more open or more back vowels.*”

¹² This modification was validated by a personal communication between Birdsong and Cooper and Ross (quoted in Birdsong, “Iconicity”, 35).

¹³ Birdsong, “Psycholinguistic Perspectives” (quoted in Birdsong, “Iconicity”, 35).

¹⁴ Ross, “Ikonismus” (quoted in Fenk-Oczlon, 521); Hagège, 184.

¹⁵ Kelly, 582. He notes that the constraint is still probabilistically respected when the second source lexeme has a higher lexical frequency than the first one.

24.9% a left source lexeme with more syllables, and 50.8% a right source lexeme with more syllables.¹⁶

2.3 Stress Alternation

The influence of a rhythmic constraint on the element ordering of binomials is mentioned in Otto Jespersen's first edition of *Growth and Structure of the English Language* in 1905.¹⁷ For Jespersen, the constraint is not dissociated from Panini's Law as he explains that the sequence *monosyllable-coordinator-disyllable* is preferred (e.g. *bread and butter* rather than *butter and bread*) because it creates a trochaic rhythm. Janet McDonald, Kathryn Bock and Michael Kelly underline that the effect of Panini's Law on element ordering needs to be distinguished from that of Frequency or Stress Alternation, and they prove experimentally that the latter constraint is valid in itself, and that, other factors being equal, it is more influential than Syllable Number.¹⁸ The stress pattern of the disyllable plays a significant role in element ordering: the disyllable is preferably first in the binomial if it is right-stressed (e.g. *surprise and sin* rather than *sin and surprise*), and second if it is left-stressed (e.g. *sin and silence* rather than *silence and sin*).¹⁹

2.4 Lexical Frequency

Włodzimierz Sobkowiak and Shih-Ping Wang have both assessed the role of lexical frequency in the element ordering of English binomials, and they conclude that the more frequent element tends to come first. Even though the lists of binomials and the frequency data are very dissimilar (Sobkowiak uses John Carrol, Peter Davies and Barry Richman's data, and Wang the British National Corpus), the results are remarkably close: the percentage of binomials respecting the constraint is respectively 63.6 and 65.5.

Frequency has also been tested positively on blends. Kelly has calculated that the source lexemes corresponding to the left splinter in a blend average 40.1 occurrences per million words whereas those represented in second position average 14.8 occurrences per million.²⁰

¹⁶ Gries, 421. Unlike Kelly, Gries includes non-coordinate items (e.g. *Chunnel, Paralympics*) in his list of blends.

¹⁷ Jespersen, 233–4; Abraham, 279.

¹⁸ McDonald, Bock and Kelly, 195–6.

¹⁹ In the experiments, the constraint is active only if the semantic constraint Animate is neutralized.

²⁰ Kelly, 582. He still obtains a statistically significant frequency differential when the source lexemes have the same number of syllables (i.e. when the role of Panini's Law is neutralized).

2.5 Prototypicality

The first reference to the prototypicality constraint dates back at least to Willi Mayerthaler, who speaks of the “*principle of perceptual accessibility*” to account for element ordering in a binomial such as *day and night*.²¹ Michael Kelly, Kathryn Bock and Franck Keil have proved experimentally that, frequency and syllable number being equal, the typicality difference between the elements of a novel binomial influences ordering: the prototypical item appears preferably before the non-prototypical item (e.g. *red* tends to occur before *gold*, *carrot* before *onion*, *biology* before *geography*,²² *football* before *sailing*). They have also shown that when a sentence containing a novel binomial is memorized and then recalled, the percentage of inversion of the non-prototype-before-prototype original order is 27 whereas it is only 13 for prototype before non-prototype. In 62% of cases, a sentence is also considered more natural if the prototype appears in first position (e.g. *the child’s errand was to buy an apple and lemon* vs. *the child’s errand was to buy a lemon and apple*).

Michael Kelly has analyzed the effect of Prototypicality on splinter order in thirty-seven coordinate blends and observes that twenty-six items obey the constraint. He underlines that the role of Prototypicality is independent of that of Frequency and of Syllable Number as the mean number of syllables is almost identical in the left and right source lexemes and he adds that the frequency differential between prototype and non-prototype in the thirty-seven blends under study is not statistically significant.²³

3. Analysis

The constraints reviewed in the previous section cannot all be experimentally tested in the same manner. Prototypicality requires complex psycholinguistic tests and measures which could not be carried out for the present study, and the remaining constraints are not all applicable to every coordinate lexical item: the phonological constraints and Stress Alternation are not applicable to blends, and Experiential Closeness, Superiority and Temporal Iconicity are only applicable to certain items of the three types of coordinate structures.

The list of coordinate lexical items under scrutiny was compiled from the combing of various present-day general English-language dictionaries and is composed of 315 binomials, 105

²¹ Mayerthaler, 13 (quoted in Edmonson, 125).

²² The category is *science*.

²³ Kelly, 583.

compounds and 142 blends. In the case of compounds and blends, for which assessing a relation of coordination is not self-evident because of the absence of an overt coordinator, the coordinate status was checked on the basis of the dictionary definitions, which had to reflect the equal semantic footing of the two elements (the two source lexemes in the case of blends). Definitional elements such as “*cross of A and B*” (e.g. for *wolf dog, broccoflower, plumcot*), “*mixture of A and B*” (e.g. for *gum resin, toxin-antitoxin, amatol, infotainment*), “*combining A and B*” (e.g. for *camiknickers, tractor-trailer, brunch, polocrosse*) and “*functioning as A and B*” (e.g. for *fighter-bomber, cafetorium, Spork*) were considered as markers of semantic equality.

A probability-based approach was adopted: constraints may be violated; predictions are statistically validated by Pearson’s chi-square test for goodness of fit, which assesses the significance of the difference between the observed data and a random distribution of the items. The expected cell counts must be equal to at least five.²⁴

3.1 Semantic Constraints

3.1.1 Experiential Closeness

Experiential Closeness predicts that the element which is empathic with the world view of the virtual prototypical speaker will come first in the coordinate lexical item. Only the most salient notions that Cooper and Ross associate with the virtual prototypical speaker – Positive, Human, Animate, Male and Adult – have been retained. The constraint can only be tested on a small number of items: two blends out of 142 (*Amerasian* and *Tex-Mex*), two compounds out of 105 (*pass-fail* and *toxin-antitoxin*) and twenty-seven binomials out of 315. Out of these twenty-seven items, twenty respect the constraint:

²⁴ Kanji, 69.

<i>Adam-and-Eve</i>	<i>do-or-die</i>	<i>make-or-break</i>
<i>all-or-nothing</i>	<i>feast-or-famine</i>	<i>old-man-and-woman</i>
<i>boom and bust</i>	<i>hen-and-chickens</i>	<i>profit and loss</i>
<i>carrot-and-stick</i>	<i>horse-and-buggy</i>	<i>sweet-and-sour</i>
<i>chicken-and-egg</i>	<i>Jekyll and Hyde</i>	<i>up-or-out</i>
<i>cops and robbers</i>	<i>life-and-death</i>	<i>wax and wane</i>
<i>cowboys and Indians</i>	<i>live-and-die</i>	

whereas seven violate it (*coach-and-four*, *dead-and-alive*, *ebb and flow*, *kill or cure*, *mom-and-pop*, *trick-or-treat* and *youth-and-old-age*). The distribution is statistically significant ($\chi^2(1) = 6.3$; $p < 0.02$). Experiential Closeness therefore proves to be an operative constraint on binomials.

3.1.2 Superiority

Superiority predicts that the element which is spatially or hierarchically superior to the other will appear first in the coordinate structure. Spatial Superiority is only applicable to four compounds (*camiknickers*, *pantyhose*, *pantywaist*²⁵ and *two-up two-down*) and four binomials (*cap and gown*, *foot and mouth*, *top-and-tail* and *ups and downs*), Hierarchical Superiority to five binomials (*chapter and verse*, *dollars-and-cents*, *five-and-ten*, *hundreds and thousands* and *nickel-and-dime*). The sample sizes of both compounds and binomials are too small to validate the prediction.

3.1.3 Temporal Iconicity

Temporal Iconicity predicts that an element referring to an event or action first in time will be ordered first in the coordinate lexical item. The constraint is applicable to seventy-two binomials (e.g. *bed and breakfast*, *born and bred*, *clean and jerk*, *cut and paste*), as well as the four blends and thirteen compounds below:

²⁵ The first two compounds conform to the constraint – “a woman’s one-piece undergarment which combines camisole and French knickers.” (*Oxford Dictionaries Online*); “a one-piece, skintight garment worn by women, combining panties and stockings.” (*Random House Unabridged Dictionary*) –, but the third one does not – “a child’s garment consisting of short pants buttoned to a waist.” (*Merriam-Webster’s Online Dictionary*).

brunch (< *breakfast* + *lunch*)

mim-mem (< *mimicry* + *memorization*)

Jacobethan (< *Jacobean* + *Elizabethan*)

twilight (< *twilight* + *night*)

cook-chill

hunter-killer

strip-search

dinner dance

murder-suicide

tie-dye

drop-kick

roll-on roll-off

washer-dryer

fly-drive

stop-go

freeze-dry

stop-start

Out of the eighty-nine items, only one, *postage and packing*, does not obey the constraint. The distribution is highly significant for both compounds ($\chi^2(1) = 13$; $p < 0.001$) and binomials ($\chi^2(1) = 68.1$; $p < 0.001$).

3.2 Phonological Constraints

3.2.1 Syllable Number

Syllable Number predicts that the number of syllables in the second element will be higher than that in the first one. The constraint is applicable to all coordinate lexical items. The left/right distribution shown in Table 1 is significant for binomials ($\chi^2(1) = 49.2$; $p < 0.001$), compounds ($\chi^2(1) = 5.8$; $p < 0.02$) and blends ($\chi^2(1) = 7.2$; $p < 0.008$), which is in accordance with Panini's Law.

Table 1 Distribution of coordinate lexical items according to the number of syllables of each element

Coordinate items	Longer left element	Longer right element
Binomials	25	105
Compounds	10	24
Blends	31	56

Because of the interaction between Syllable Number and Frequency, neutralizing the role of the latter may shed light on the exact influence of the former on the three types of coordinate items. The three lists were therefore reduced to only those lexical items containing a right element which is more frequent than the left one.²⁶ The left/right distribution of non-equisyllabic binomials²⁷ shown

²⁶ For information on the frequency data used, see below (3.4).

in Table 2 is still statistically significant ($\chi^2(1) = 6.4$; $p < 0.02$), whereas the results are no longer significant for compounds and blends, which indicates that, for the last two types, the influence of Syllable Number on element ordering is not independent of that of Frequency.

Table 2 Distribution of coordinate lexical items with a more frequent right element according to the number of syllables of each element

Coordinate items	Longer left element	Longer right element
Binomials	12	28
Compounds	4	5
Blends	22	19

3.2.2 Onset Constraints

Initial Consonant Complexity predicts that the number of initial consonants will be higher in the second element of the coordinate lexical item than in the first one. The constraint is applicable to equisyllabic compounds and binomials and to a type of blends. The left/right distribution shown in Table 3 is significant neither for compounds nor for binomials.

Table 3 Distribution of equisyllabic coordinate lexical items according to the degree of complexity of the onset of each element

Equisyllabic items	More complex left onset	More complex right onset
Compounds	16	10
Binomials	31	27

If the analysis is limited to those items whose elements have the same superrime (i.e. which are identical except for their onset),²⁸ three compounds and nineteen binomials remain:

²⁷ For brevity's sake, the adjectives *equisyllabic* and *non-equisyllabic* are sometimes directly used with *lexical item*, *binomial*, *compound* and *blend* to refer to items whose elements are equisyllabic or non-equisyllabic.

²⁸ The term is borrowed from Thomas Berg.

happy-clappy

teeny-weeny

tie-dye

chalk and talk

meet and greet

surf and turf

chew and spew

name and shame

thrills and spills

cut and shut

odds and sods

wear and tear

high and dry

pan and scan

wheel and deal

huff and puff

pump and dump

wine and dine

hustle and bustle

run and gun

make or break

shake and bake

The number of initial consonants is identical in fifteen cases, and for the remaining seven items (*happy-clappy*, *chew and spew*, *high and dry*, *make or break*, *meet and greet*, *odds and sods* and *pan and scan*), the constraint is systematically respected. The sample size is however too low for the result to have statistical significance.

Blends behave opposite to the prediction. With the exception of *mingy* (< *mean* + *stingy*), blends which originate from a source lexeme with a simple onset and another with a complex onset and whose source lexemes are both clipped at the onset-superrime breakpoint all retain the complex onset:

breakfast + *lunch* > *brunch* (~~leak~~fast)

smoke + *haze* > *smaze* (~~hoke~~)

fries + *rings* > *frings* (~~ries~~)

smoke + *fog* > *smog* (~~foke~~)

fry + *sizzle* > *frizzle* (~~sy~~)

spoon + *fork* > *Spork* (~~Foon~~)

pliers + *wrench* > *plench* (~~wriers~~)

squirm + *wiggle* > *squiggle* (~~wirm~~)

prim + *sissy* > *prissy* (~~sim~~)

twist + *fiddle* > *twiddle* (~~fist~~)

skirt + *short* > *skort* (~~shirt~~)

The distribution is statistically significant ($\chi^2(1) = 8.3$; $p < 0.004$). It is therefore the source lexeme with the higher number of initial consonants which is expected to come first.

Initial Consonant Obstruency predicts that the initial consonant of the left element will be less obstruent than that of the right element. The constraint is applicable to equisyllabic compounds and binomials whose elements have simple onsets. The obstruency scale used for the analysis consists of five levels: stops > fricatives > nasals > liquids > glides. The left/right distribution shown in Table 4 is significant for binomials ($\chi^2(1) = 6.2$; $p < 0.02$), but not for compounds.

Table 4 Distribution of equisyllabic coordinate lexical items with simple-onset elements according to the degree of obstruency of each onset

Equisyllabic items	More obstruent left onset	More obstruent right onset
Compounds	9	8
Binomials	25	46

If the analysis is restricted to those items whose elements have the same superrime and a simple onset, only two compounds (*teeny-weeny*, *tie-dye*) and eleven binomials remain. Nine binomials (*huff and puff*, *hustle and bustle*, *name and shame*, *run and gun*, *shake and bake*, *surf and turf*, *wear and tear*, *wheel and deal* and *wine and dine*) obey the constraint, one violates it (*cut and shut*), and the onsets of the elements of the last binomial (*pump and dump*) have the same degree of obstruency; the left/right distribution is statistically significant ($\chi^2 (1) = 6.4$; $p < 0.02$), which confirms the validity of Initial Consonant Obstruency for binomials.

3.2.3 Nucleus Constraints

Vowel Length predicts that the right element will contain a longer vowel than the left element. The constraint is applicable to all compounds and binomials composed of two monosyllables. The left/right distribution of compounds shown in Table 5 is not significant, whereas that of binomials is highly significant ($\chi^2 (1) = 38.5$; $p < 0.001$).

Table 5 Distribution of coordinate lexical items according to the nucleus length of each element

Coordinate items	Longer left nucleus	Longer right nucleus
Compounds	4	9
Binomials	10	63

Vowel Quality predicts that the right element will contain a more open and/or more back vowel than the left element. The constraint is applicable to non-blended lexical items composed of two monosyllables which both contain a monophthong. The left/right distribution shown in Table 6 is significant neither for compounds nor for binomials.

Table 6 Distribution of coordinate lexical items according to the degree of backness and/or openness of the nucleus of each element

Coordinate items	More back and/or open left nucleus	More back and/or open right nucleus
Compounds	4	6
Binomials	22	27

3.2.4 Coda Constraints

Final Consonant Complexity predicts that the right element will contain more final consonants than the left element. The constraint is applicable to all equisyllabic compounds and binomials. The left/right distribution shown in Table 7 is not significant either for compounds or for binomials.

Table 7 Distribution of equisyllabic coordinate lexical items according to the degree of complexity of the coda of each element

Equisyllabic items	More complex left coda	More complex right coda
Compounds	8	7
Binomials	34	34

Final Consonant Obstruency predicts that the element containing a more obstruent final consonant will appear first in the coordinate lexical item. The constraint is applicable to all equisyllabic compounds and binomials whose elements have simple codas. The left/right distribution shown in Table 8 is significant for binomials ($\chi^2 (1) = 4.4$; $p < 0.04$), but not for compounds.

Table 8 Distribution of equisyllabic coordinate lexical items with simple-coda elements according to the degree of obstruency of each coda

Equisyllabic items	More obstruent left coda	More obstruent right coda
Compounds	9	4
Binomials	37	21

3.3 Stress Alternation

Stress Alternation predicts that element ordering will respect a strict alternation between stressed and unstressed syllables. The constraint is applicable to binomials which combine a monosyllable and a disyllable. Out of ninety-four items, sixty-seven have a 2010²⁹ stress pattern (e.g. *bits and pieces, noughts and crosses*), one has a 0201 pattern (*alive and well*), and twenty-six have a 2001 pattern (e.g. *pay and display, rhythm and blues*). The discrepancy between the two patterns respecting Stress Alternation and the pattern violating it is statistically highly significant ($\chi^2 (1) = 18.8$; $p < 0.001$). The analysis can then be narrowed to only those compounds for which Temporal

²⁹ “1” symbolizes a syllable bearing primary stress, “2” a syllable bearing secondary stress, and “0” an unstressed syllable.

Iconicity, which is almost never violated, is not applicable. Out of the remaining seventy-eight items, fifty-four respect Stress Alternation ($\chi^2(1) = 11.5$; $p < 0.001$), which confirms the validity of the prediction. Besides, it may be enlightening to distinguish between the effect of Stress Alternation and that of Syllable Number on element ordering. As the vast majority of English disyllables are left-stressed – Cynthia Clopper has calculated that 78.5% of the 4619 disyllables listed in the *Webster's Pocket Dictionary* are so³⁰ – and tetrasyllabic binomials overwhelmingly respect Syllable Number, most of these binomials are expected to have a 2010 stress pattern. As a consequence, the list of binomials under scrutiny needs to be limited to tetrasyllabic items containing a right-stressed disyllable. If the lexical items whose element ordering is semantically constrained are discarded, only two binomials remain (*aid and abet* and *dead-and-alive*), which makes any statistical extrapolation impossible. Stress Alternation is therefore validated for binomials, but its role cannot be clearly differentiated from that of Syllable Number. As for compounds, only ten items are relevant. The distribution is not significant as six obey the constraint (*bittersweet, dinner dance, hammer-axe, pantyhose, pantywaist* and *sofa bed*) whereas four violate it (*fridge freezer, gum resin, queen mother* and *work-study*).

3.4 Lexical Frequency

Frequency predicts that the more frequent element will appear in first position. The constraint is applicable to all coordinate lexical items. The frequency data were taken from Adam Kilgarriff's frequency lists; only those lemmatized word-forms which appear at least five times in the British National Corpus were retained. The left/right distribution shown in Table 9 is significant neither for compounds nor for blends, but it is significant for binomials ($\chi^2(1) = 6.3$; $p < 0.02$).

Table 9 Distribution of coordinate lexical items according to the frequency of their elements

Coordinate items	More frequent left element	More frequent right element
Binomials	169	126
Compounds	40	34
Blends	59	70

Because of the correlation between Panini's Law and Frequency (more frequent words tend to be shorter), the role of Syllable Number was then neutralized. As none of the distributions shown in Table 10 differs statistically from the null hypothesis of random ordering, it can be concluded that,

³⁰ Clopper, 5.

when significant, the influence of Frequency on the element ordering of coordinate lexical items is not independent of that of Syllable Number.

Table 10 Distribution of equisyllabic coordinate lexical items according to the frequency of their elements

Equisyllabic items	More frequent left element	More frequent right element
Binomials	88	86
Compounds	21	23
Blends	20	29

4. General perspective

Several abovementioned constraints have been claimed to be operative in a variety of typologically diverse languages.³¹ In Basque, William Jacobsen holds that coordinate compounds respect Experiential Closeness, Superiority, Temporal Iconicity and Syllable Number. In Polish, Jolanta Szpyra makes the case that binomials obey the same general semantic and phonological principles as in English. In Hungarian, László Pordány maintains that both binomials and compounds respect the semantic constraints that are operative in English as well as Syllable Number, Vowel Quality and Initial Consonant Complexity. In Hindi, Shaligram Shukla contends that Experiential Closeness, Superiority, Syllable Number and Initial Consonant Complexity are operative on coordinate compounds.³² In Japanese, Taro Kageyama claims that coordinate compounds obey Experiential Closeness and Laurence Labrune affirms that they respect Initial Consonant Obstruency.³³

These data lead to postulate that general cognitive principles might be at work in the element ordering of lexical coordinate items. For Experiential Closeness and Superiority, the explanation probably lies in the existence of cognitive universals, and for Temporal Iconicity, it is the iconic nature of the constraint which makes it valid in virtually all languages. As for phonological constraints, Cooper and Ross, Birdsong and Allan all stress that they conspire to magnify the opposition between the first element – which is structurally lighter, shorter, simpler – and the second – which, in contrast, is heavier, longer, more complex.³⁴ A cognitivist explanation for the element ordering of coordinate lexical items put forward by Cooper and Ross, Sobkowiak, Birdsong

³¹ It is to be noted that, with the exception of Labrune, the various authors do not support their claims with tests of statistical significance.

³² Shukla, 108–9.

³³ Kageyama, 237; Labrune, 18–24.

³⁴ Cooper and Ross, 80; Birdsong, “Psycholinguistic Perspectives” (quoted in Birdsong, “Iconicity”, 35).

and William Cooper and Gayle Klouda goes one step further by linking the semantic and phonological constraints: both tend to place in first position the element which is informationally poorer and therefore cognitively easier to process.³⁵ Information processing is optimized if speech structure follows the principle demanding that information be distributed asymmetrically, in order to alternate lighter and heavier elements. This principle is not limited to the element ordering of coordinate structures; it is pervasive, and applicable to the whole of grammar, as Kathryn Bock and John Hawkins³⁶ have both demonstrated in their studies of linguistic structure and information processing.

5. Conclusion

Twelve constraints which are said to influence the element ordering of coordinate lexical items were tested on 562 English binomials, compounds and blends, and for each constraint statistical significance was assessed. The main finding of the study is that a majority of constraints are operative, which was not necessarily expected as many authors formulate constraints on the basis of a small number of items and do not validate their claims with tests of statistical significance. As summarized in Table 11, Syllable Number is validated for the three types of coordinate lexical items, and Temporal Iconicity for both binomials and compounds. In addition, the element ordering of binomials has been proved to be influenced by Experiential Closeness, Vowel Length (for items combining two monosyllables), Initial Consonant Obstruency (for equisyllabic items composed of two simple-onset elements), Final Consonant Obstruency (for equisyllabic items composed of two simple-coda elements), Stress Alternation (for items composed of a monosyllable and a disyllable) and Lexical Frequency, and that of blends by Initial Consonant Complexity (for items originating from a source lexeme with a simple onset and another with a complex onset and whose source lexemes are both clipped at the onset-superrime breakpoint).

³⁵ Cooper and Ross, 92; Sobkowiak, 412; Birdsong, “Iconicity”, 35; Cooper and Klouda, 339. Frequency should be added as psycholinguists have amply demonstrated that the accessibility of a word correlates with its frequency (Babin, 20–4; Le Ny, 144–7).

³⁶ Hawkins, *Performance Theory* and *Efficiency and Complexity*.

Table 11 Operative constraints on the element ordering of English coordinate lexical items

Constraint	Binomials	Compounds	Blends
Experiential closeness	✓		
Temporal iconicity	✓	✓	
Syllable number	✓	✓	✓
Initial consonant complexity			✓
Initial consonant obstruency	✓		
Vowel length	✓		
Final consonant obstruency	✓		
Stress alternation	✓		
Lexical frequency	✓		

Acknowledgments

I am indebted to Anne Viallefont for help with statistics, and grateful to Nicolas Ballier and anonymous reviewers for comments and suggestions on earlier drafts of this paper. The usual disclaimers apply.

References

- Abraham, Richard D. "Fixed Order of Coordinates: A Study in Lexicography". *Modern Language Journal* 34, no. 4 (1950): 276–87.
- Allan, Keith. "Hierarchies and the Choice of Left Conjuncts (with Particular Attention to English)". *Journal of Linguistics* 23, no. 1 (1987): 51–77.
- Arnaud, Pierre J. L. "Problématique du nom composé". In *Le Nom composé: Données sur seize langues*, ed. Pierre J. L. Arnaud, 329–53. Lyon: Presses Universitaires de Lyon, 2004.
- Babin, Jean-Philippe. *Lexique mental et morphologie lexicale*. Bern: Lang, 1998.
- Berg, Thomas. "On the Internal Structure of Polysyllabic Monomorphemic Words: The Case for Superrimes". *Studia Linguistica* 43, no. 1 (1989): 5–32.
- Birdsong, David. "Psycholinguistic Perspectives on the Phonology of Frozen Word Order". PhD diss., Harvard University, 1979.
- Birdsong, David. "Iconicity, Markedness, and Processing Constraints in Frozen Locutions". In *Syntactic Iconicity and Linguistic Freezes: The Human Dimension*, ed. Marge Landsberg, 31–45. Berlin: De Gruyter, 1995.

- Bock, J. Kathryn. "Toward a Cognitive Psychology of Syntax: Information Processing Contributions to Sentence Formulation". *Psychological Review* 89, no. 1 (1982): 1–47.
- Carroll, John B., Peter Davies and Barry Richman. *The American Heritage Word Frequency Book*. Boston: Houghton Mifflin, 1971.
- Clopper, Cynthia G. "Frequency of Stress Patterns in English: A Computational Analysis". *Indiana University Linguistics Club Working Papers Online* 2 (2002): <https://www.indiana.edu/~iulcwp/pdfs/02-clopper02.pdf>.
- Cooper, William E. and Gayle V. Klouda. "The Psychological Basis of Syntactic Iconicity". In *Syntactic Iconicity and Linguistic Freezes: The Human Dimension*, ed. Marge Landsberg, 331–41. Berlin: De Gruyter, 1995.
- Cooper, William E. and John R. Ross. "World Order". In *Papers from the Parasession on Functionalism; April 17, 1975*, ed. Robin E. Grossman, L. James San and Timothy J. Vance, 63–111. Chicago: Chicago Linguistic Society, 1975.
- Edmonson, Jerry A. "Biological Foundations of Language Universals". In *Developmental Mechanisms of Language*, ed. Charles-James N. Bailey and Roy Harris, 109–30. Oxford: Pergamon, 1985.
- Ertel, Suitbert. "Where do the Subjects of Sentences Come from?" In *Sentence Production: Developments in Research and Theory*, ed. Sheldon Rosenberg, 141–68. Hillsdale, NJ: Erlbaum, 1977.
- Fenk-Oczlon, Gertraud. "Word Frequency and Word Order in Freezes". *Linguistics* 27, no. 3 (1989): 517–56.
- Gries, Stefan Th. "Isn't That Fantabulous? How Similarity Motivates Intentional Morphological Blends in English". In *Language, Culture, and Mind*, ed. Michel Achard and Suzanne Kemmer, 415–28. Stanford, CA: CSLI Publications, 2004.
- Hagège, Claude. *L'Homme de paroles: Contribution linguistique aux sciences humaines*. Paris: Fayard, 1985.
- Hawkins, John A. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press, 1994.
- Hawkins, John A. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press, 2004.
- Jacobsen, William H. "Basque Copulative Compounds: A Problem in Irreversible Binomials". In *Proceedings of the Eighth Annual Meeting of the Berkeley Linguistics Society*, ed. M. Macaulay, O. D. Gensler, C. Brugman, I. Civkulus, A. Dahlstrom, K. Krile and R. Sturm, 384–97. Berkeley, CA: Berkeley Linguistics Society, 1982.
- Jespersen, Otto. *Growth and Structure of the English Language*. Leipzig: Teubner, 1905.
- Kageyama, Taro. "Word Formation in Japanese". *Lingua* 57, no. 2–4 (1982): 215–58.
- Kangi, Gopal K. *100 Statistical Tests, New Edition*. London: Sage, 1999.
- Kelly, Michael H. "To 'brunch' or to 'brench': Some Aspects of Blend Structure". *Linguistics* 36, no. 3 (1998): 579–90.
- Kelly, Michael H., J. Kathryn Bock and Franck C. Keil. "Prototypicality in a Linguistic Context: Effects on Sentence Structure". *Journal of Memory and Language* 25, no. 1 (1986): 59–74.

- Kilgarriff, Adam. "BNC Database and Word Frequency Lists". <http://www.kilgarriff.co.uk/bnc-readme.html>.
- Kuno, Susumu. "On the Interaction between Syntactic Rules and Discourse Principles". In *Explorations in Linguistics: Papers in Honor of Kazuko Inoue*, ed. George Bedell, Eichi Kobayashi and Masatake Muraki, 279–304. Tokyo: Kenkyusha, 1979.
- Labrune, Laurence. "Patterns of Phonemic Preferences in Japanese Non-Headed Binary Compounds: What *waa-puro*, *are-kore* and *mecha-kucha* Have in Common". *Gengo Kenkyu* 129 (2006): 3–41.
- Landsberg, Marge E. "Semantic Constraints on Phonologically Independent Freezes". In *Syntactic Iconicity and Linguistic Freezes: The Human Dimension*, ed. Marge Landsberg, 65–78. Berlin: De Gruyter, 1995.
- Le Ny, Jean-François. *Comment l'esprit produit du sens: Notions et résultats des sciences cognitives*. Paris: Odile Jacob, 2005.
- Malkiel, Yakov. "Studies in Irreversible Binomials". *Lingua* 8 (1959): 113–60.
- Mayerthaler, Willi. *Morphologische Natürlichkeit*. Wiesbaden: Athenaion, 1981.
- McDonald, Janet L., J. Kathryn Bock and Michael H. Kelly. "Word and World Order: Semantic, Phonological, and Metrical Determinants of Serial Position". *Cognitive Psychology* 25, no. 2 (1993): 188–230.
- Merriam-Webster's Online Dictionary*. <http://www.merriam-webster.com/dictionary>.
- Oakeshott-Taylor, John. "Phonetic Factors in Word Order". *Phonetica* 41, no. 4 (1984): 226–37.
- Oxford Dictionaries Online*. <http://oxforddictionaries.com>.
- Picone, Michael D. *Anglicisms, Neologisms and Dynamic French*. Amsterdam: Benjamins, 1996.
- Pinker, Steven and David Birdsong. "Speakers' Sensitivity to Rules of Frozen Word Order". *Journal of Verbal Learning and Verbal Behavior* 18, no. 4 (1979): 497–508.
- Pordány, László. "Irreversible Binomials in Hungarian". *Acta Linguistica Academiae Scientiarum Hungaricae* 36 (1986): 151–79.
- Random House Unabridged Dictionary*. <http://dictionary.infoplease.com>.
- Ross, John R. "Myopia". Paper presented at the Fifth Conference on New Ways of Analyzing Variation in English, Georgetown University, Washington, D. C., October 1976.
- Ross, John R. "Ikonismus in der Phraseologie". *Zeitschrift für Semiotik* 2 (1980): 39–56.
- Shukla, Shaligram. *Hindi Morphology*. Munich: Lincom Europa, 2001.
- Sobkowiak, Włodzimierz. "Unmarked-before-marked as a Freezing Principle". *Language and Speech* 36, no. 4 (1993): 393–414.
- Szpyra, Jolanta. "Semantic and Phonological Constraints on Conjunct Ordering in English and Polish". *Kwartalnik Neofilologiczny* 30, no. 1 (1983), 33–53.
- Van Langendonck, Willy. "Categories of Word Order Iconicity". In *Syntactic Iconicity and Linguistic Freezes: The Human Dimension*, ed. Marge Landsberg, 79–90. Berlin: De Gruyter, 1995.
- Wang, Shih-Ping. "Mutual Information and Corpus-based Approaches to Reduplicative Fixed Expressions". In *Proceedings of the Corpus Linguistics 2003 Conference* (= Technical Papers, Vol. 16), ed. Dawn Archer, Paul Rayson, Andrew Wilson and Tony McEnery, 869–78. Lancaster: University Centre for Computer Corpus Research on Language, 2003.